



## Relation between NOD2 genotype and changes in innate signaling in Crohn's disease on mRNA and miRNA levels

Chen, Yun; Salem, Mohammad; Boyd, Mette; Bornholdt, Jette; Li, Yuan; Coskun, Mehmet; Seidelin, Jakob Benedict; Sandelin, Albin; Nielsen, Ole Haagen

*Published in:*  
npj Genomic Medicine

*DOI:*  
[10.1038/s41525-016-0001-4](https://doi.org/10.1038/s41525-016-0001-4)

*Publication date:*  
2017

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)

*Citation for published version (APA):*  
Chen, Y., Salem, M., Boyd, M., Bornholdt, J., Li, Y., Coskun, M., ... Nielsen, O. H. (2017). Relation between NOD2 genotype and changes in innate signaling in Crohn's disease on mRNA and miRNA levels. *npj Genomic Medicine*, 2, [3]. <https://doi.org/10.1038/s41525-016-0001-4>

## ARTICLE OPEN

Relation between *NOD2* genotype and changes in innate signaling in Crohn's disease on mRNA and miRNA levels

Yun Chen<sup>1</sup>, Mohammad Salem<sup>2</sup>, Mette Boyd<sup>1</sup>, Jette Bornholdt<sup>1</sup>, Yuan Li<sup>2</sup>, Mehmet Coskun<sup>1,2</sup>, Jakob Benedict Seidelin<sup>2</sup>, Albin Sandelin<sup>1</sup> and Ole Haagen Nielsen<sup>2</sup>

Crohn's disease is associated with an altered innate immune response of pathogenic importance. This altered response can be associated to loss-of-function polymorphisms in the *NOD2* (nucleotide-binding oligomerization domain-containing protein 2) gene, but also changes in transcriptional and post-transcriptional regulatory layers, including microRNA activity. Here, we characterized the link between *NOD2* genotype and inflammatory-mediated changes in innate signaling by studying transcriptional and post-transcriptional activity in response to *NOD2*-agonist muramyl dipeptide in monocytes from healthy controls, and Crohn's disease patients with and without *NOD2* loss-of-function polymorphisms. We measured the expression of genes and microRNAs in monocytes from these subjects after stimulation with muramyl dipeptide. Gene expression profiles mainly distinguished the actual muramyl dipeptide response, but not the genotype. A hyper-responsive phenotype was found in Crohn's disease patients without *NOD2* mutations, characterized by upregulated cytokine receptors and general downregulation of microRNA expression. Conversely, microRNA expression could identify genotype-specific differences between subject groups but exhibited little change upon muramyl dipeptide treatment. Only two microRNAs showed muramyl dipeptide-induced response, including miR-155, which was found to regulate multiple genes and whose host gene was one of the highest muramyl dipeptide responders. miR-155 was upregulated in Crohn's disease patients with *NOD2* mutations following lipopolysaccharide and *Escherichia coli* treatment, but the upregulation was substantially reduced upon muramyl dipeptide treatment. While Crohn's disease patients with *NOD2* mutations on average showed a reduced muramyl dipeptide response, the cohort exhibited large individual variance: a small subset had inflammatory responses almost comparable to wild-type patients on both gene and miR-155 regulatory levels.

npj Genomic Medicine (2017)2:3; doi:10.1038/s41525-016-0001-4

## INTRODUCTION

Crohn's disease (CD)<sup>1</sup> constitutes together with ulcerative colitis (UC)<sup>2</sup> the two most prevailing disorders under the umbrella term inflammatory bowel disease (IBD).<sup>3</sup> CD can affect any part of the gastrointestinal tract but most often the ileocecal region. General clinical manifestations of intermitting bowel inflammation are associated with abdominal pain, fever, prolonged diarrhea, and/or weight loss.<sup>4</sup> In CD, genetic variations and environmental factors such as altered microbiota interact to produce the inflammatory background of the disease.<sup>1, 5</sup> Along with changes in the adoptive immune system, an impaired innate immunity is believed to play a crucial role in the immunopathogenesis of CD.

Epidemiologic and linkage studies suggest that genetic factors play a clinical role in determining CD susceptibility.<sup>6, 7</sup> One of the key genes identified through these studies is *NOD2* (nucleotide-binding oligomerization domain-containing protein 2),<sup>8, 9</sup> encoding a cytosolic pathogen recognition receptor that specifically binds to muramyl dipeptide (MDP), the smallest bioactive component of peptidoglycans present in most bacteria.<sup>10</sup> Activation of *NOD2* leads to induction of key immune signaling pathways involved in CD pathogenesis, including the upregulation of various pro-inflammatory cytokines.<sup>11</sup> In addition, continuous MDP stimulation inhibits the development of experimental colitis

in animal models,<sup>12</sup> suggesting that normally functioning *NOD2* signaling pathways might be directly involved in protecting and maintaining the intestinal homeostasis. Furthermore, innate signaling is altered even in *NOD2* wild-type (WT) CD patients,<sup>13</sup> implying post-transcriptional modifications of *NOD2* signaling pathways could be involved in re-establishment of normal innate responses, and this mechanism could be a future therapeutic target aiming at restoring microbial responses in CD. In this context, recent studies have linked *NOD2* to crucial post-transcriptional changes, including regulation by microRNAs (miRNAs) in different intestinal epithelial cell lines.<sup>14–16</sup> miRNAs are short structured RNAs that play important roles in the regulation of gene transcription and translation in development and disease progression, including CD.<sup>17, 18</sup> One of the most well-known miRNA regulators associated with inflammation and innate immunity is miR-155.<sup>19–22</sup> Previous studies have shown an elevated level of miR-155 in colonic mucosa of CD and UC,<sup>23, 24</sup> and depletion of miR-155 protects from experimental colitis in mice.<sup>25</sup> Different pathogenic stimuli induce miR-155 in vitro, including viral infection<sup>20, 26, 27</sup> and lipopolysaccharide (LPS).<sup>28</sup> However, the role of miR-155 in immune cells in the context of CD and *NOD2* polymorphisms is not well established.

<sup>1</sup>The Bioinformatics Centre, Department of Biology and Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen, Denmark and <sup>2</sup>Department of Gastroenterology, Medical Section, Herlev Hospital, University of Copenhagen, Herlev Ringvej 75, DK-2730 Herlev, Denmark  
Correspondence: Jakob Benedict Seidelin (jseidelin@dadlnet.dk) or Albin Sandelin (albin@binf.ku.dk) or Ole Haagen Nielsen (ole.haagen.nielsen@regionh.dk)

Received: 13 May 2016 Revised: 11 November 2016 Accepted: 11 November 2016

Published online: 08 February 2017

There are three major single-nucleotide polymorphisms (SNPs) of the *NOD2* gene (SNP8, SNP12, and SNP13) that are strongly associated to CD and cause a loss-of-function phenotype with a reduced response to MDP.<sup>11</sup> SNP8 and SNP12 are missense mutations, whereas SNP13 is a frameshift mutation with the most drastic loss-of-function phenotype.<sup>29</sup> The link between SNPs in *NOD2* and CD-associated inflammatory responses might have implications for the disease not only because an innate immune dysfunction is important for the pathogenesis of CD, but also since normalization of innate signaling alleviates colitis.<sup>12, 30</sup> However, controversies exist on how alterations in *NOD2* signaling can lead to an increased susceptibility to inflammation, and what the role of post-transcriptional regulation in patients with/without disease-associated *NOD2* variants are for the resulting pathogenic innate immune response.<sup>11</sup> Therefore, transcriptional profiles of *NOD2* mutant cells responding to MDP or bacterial stimulation have been analyzed using coverage-limited microarrays and focused either on mRNAs or miRNAs. The miRNA-mRNA regulation studies were either a single-target study<sup>31</sup> or done in non-primary cells or immune cells derived from non-CD subjects.<sup>15, 32</sup> Thus, it is important to investigate the interactions between miRNAs and genes genome-wide in patient-derived cells, and study how this is affected in innate immune response as function of different *NOD2* genotypes. To do this, it is necessary to investigate both RNA species from the same cells upon inflammation induction in cells with and without *NOD2* loss-of-function SNPs.

To this end, in this study we measured the expression of genes and miRNAs in human blood monocytes before and after MDP stimulation using genome-wide RNA and small RNA sequencing in an exploratory cohort consisting of healthy control subjects and CD patients with/without *NOD2* loss-of-function SNPs ( $N=7$  in total), and subsequently validated our major findings in larger patient groups ( $N=29$ , in total) as well as all identified miRNA-gene interaction candidates in the THP-1 cell line ( $N=8$ ). Our data suggest that gene expression effectively identifies pathological MDP/*NOD2*-associated immune gene responses in blood monocytes from subjects without *NOD2* loss-of-function mutations. A pre-inflammatory, hyper-responsive phenotype was found in CD patients without *NOD2* loss-of-function SNPs. Conversely, the miRNA regulatory layer was surprisingly static before and after MDP stimulation, but could predict subject and genotype groups regardless of MDP treatment. Only two miRNAs, miR-190A and miR-155, responded to MDP. CD patients with *NOD2* loss-of-function SNPs were, as expected, on average less responsive to MDP compared to other subjects at both gene and miR-155 levels, but these patients displayed a surprisingly large variance in their responses ranging from a total loss to a nearly intact inflammatory response. Although *NOD2* loss-of-function SNP patients displayed reduced miR-155 induction after MDP treatment, miR-155 induction was unchanged when treating with LPS or intact bacteria treatment, suggesting that *NOD2* mutations only affect *NOD2*-pathway-elicited response and not other inflammatory pathways.

## RESULTS

Differential expression profiles between CD<sub>WT</sub>, CD<sub>NOD2</sub>, and control subjects

We first analyzed the miRNA and gene expression signatures between the three subject groups: CD<sub>WT</sub>, CD<sub>NOD2</sub>, and controls using small RNA-Seq and RNA-Seq as defined in the "Methods" section before MDP treatment (green arrows, Fig. 1 and top of Fig. 2a,b) and after MDP treatment (yellow arrows, Fig. 1 and top of Fig. 2c,d).

Intriguingly, the miRNA expression signature could separate the three subject groups (controls, CD<sub>NOD2</sub>, and CD<sub>WT</sub>) both before and after MDP treatment. In both states, control subjects and CD<sub>NOD2</sub> patients were more similar to each other than CD<sub>WT</sub> patients, where many miRNAs were more lowly expressed

compared to the other two groups (Fig. 2a,c). Thus, the overall miRNA population was downregulated in CD<sub>WT</sub> patients, suggesting an overall lower impact of miRNA-mediated gene repression in this group.

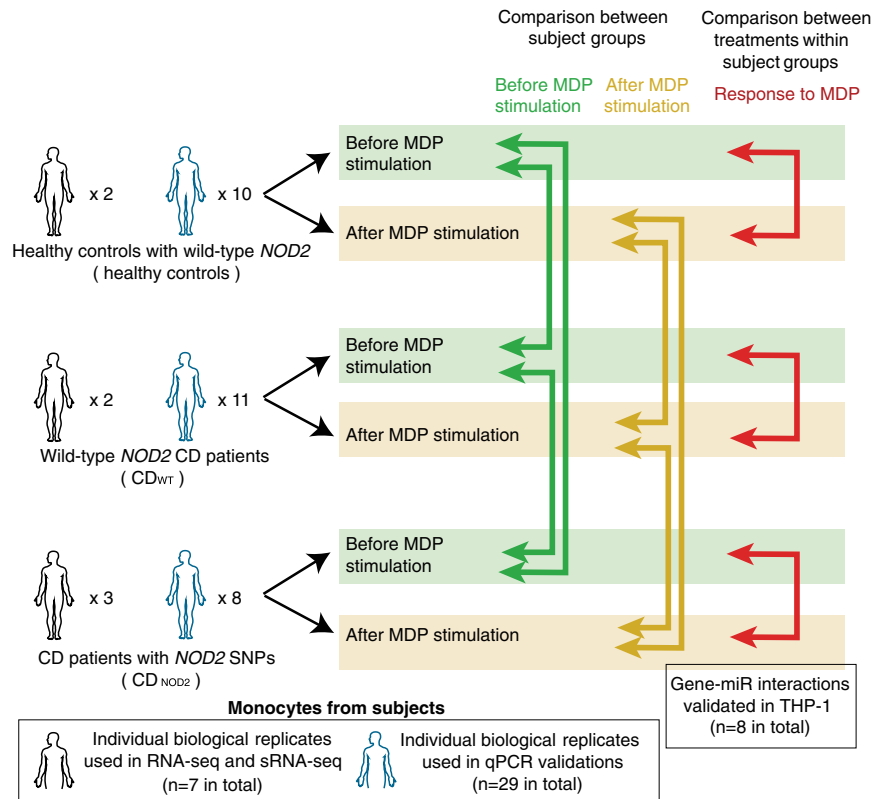
In contrast, the differences of gene expression could not distinguish subject groups as clearly as the miRNA expression, because the differences between subject groups were much smaller than that observed by miRNA expression (see dendrogram in Fig. 2b,d, and Supplementary Fig. 1a, b).

We next analyzed the differentially expressed miRNAs and genes identified in these two analyses. Nearly half of miRNAs (42%, 5/12) that were differentially expressed between subject groups before MDP treatment were also differentially expressed between subject groups after MDP treatment (Fig. 2e, Venn diagram in the left panel). The expression levels of these shared miRNAs were highly consistent before and after MDP treatment (Supplementary Fig. 2a); thus, the overall regulatory layer of miRNAs seemed to be independent of MDP signaling between groups, and could at the same time separate the subject groups. Indeed, the five miRNAs that were differentially expressed between groups regardless of treatment could alone distinguish subject groups (but not MDP treatment) efficiently (Fig. 2e, dendrogram in the right panel).

In contrast, only around 10% (13/108, Fig. 2f, also see Supplementary Fig. 1c) of genes that were significantly different between subject groups before MDP treatment were also significantly different between patient groups after MDP treatment. These shared genes could not separate the subject groups (Fig. 2f, dendrogram in the left panel). Consistently, these genes displayed a wide expression difference between subjects before and after MDP (Supplementary Fig. 2b). Genes that were differentially expressed between subject groups before MDP stimulation showed an enrichment of gene ontology (GO) terms relating to defense response, inflammatory response, response to wounding, and response to virus (Fig. 3a and Supplementary Fig. 3a). In contrast, genes differentially expressed between patient genotype groups after MDP treatment were enriched in chemokine and cytokine-related functions (Fig. 3b and Supplementary Fig. 3b). Quantitative PCR (qPCR) analysis for two selected chemokines, *CXCL1* and *CCL2*, in a larger cohort confirmed upregulation after MDP treatment in healthy controls and CD<sub>WT</sub>, but not in CD<sub>NOD2</sub> patients (Fig. 3c,d; Supplementary Fig. 3c, d show corresponding RNA-Seq data). This is consistent with the established link between *NOD2* and cytokine induction through the NF- $\kappa$ B pathway.<sup>11</sup>

miRNA and gene expression response to MDP

In order to investigate cell response to MDP, we compared the difference in miRNA and gene expression before vs. after MDP treatment within each patient group (red arrows in Fig. 1b and top of Fig. 4). Consistent with the analysis above, only two miRNAs (miR-155 and miR-190A) were significantly changed as a response to MDP in any pairwise comparison. Because miR-155 was much higher expressed than miR-190A (Fig. 4a, left panel), and was previously established as a key inflammatory regulator,<sup>21, 22, 33</sup> we validated its expression change by qPCR in the larger validation cohort. The results showed that miR-155 was significantly upregulated in controls and CD<sub>WT</sub>, but not in CD<sub>NOD2</sub> after MDP treatment ( $P \leq 0.05$ , paired two-sided *t*-test) (Fig. 4a, right panel; Supplementary Fig. 4a shows corresponding small RNA-Seq data). Notably, there was no difference in miR-155 upregulation between CD<sub>WT</sub> and CD<sub>NOD2</sub> patients when cells were treated with either LPS or intact intestinal *Escherichia coli* bacteria (Fig. 4b). Thus, miR-155 could be upregulated by multiple inflammatory stimuli but only the MDP-triggered upregulation was dependent on the *NOD2* genotype. This makes sense, as *NOD2* is a specific sensor for MDP,



**Fig. 1** Overview of cohorts and differential expression analyses. Monocytes were derived from three subject groups: healthy controls, CD patients with WT *NOD2* genotypes ( $CD_{WT}$ ), and CD patients with loss-of-function *NOD2* genotypes ( $CD_{NOD2}$ ). Monocytes were analyzed by RNA-Seq and small RNA-Seq (black body outlines, number of tested individuals is indicated), and selected genes and miRNAs were validated by qPCR in a validation cohort (blue body outlines, number of tested individuals is indicated). We analyzed the difference between subject groups, before and after MDP stimulation (green and yellow arrows, shown in Figs. 2,3), and the response to MDP stimulation within each group (red arrows, shown in Fig. 4)

while LPS and *E. coli* activate other *NOD2*-independent inflammatory pathways.

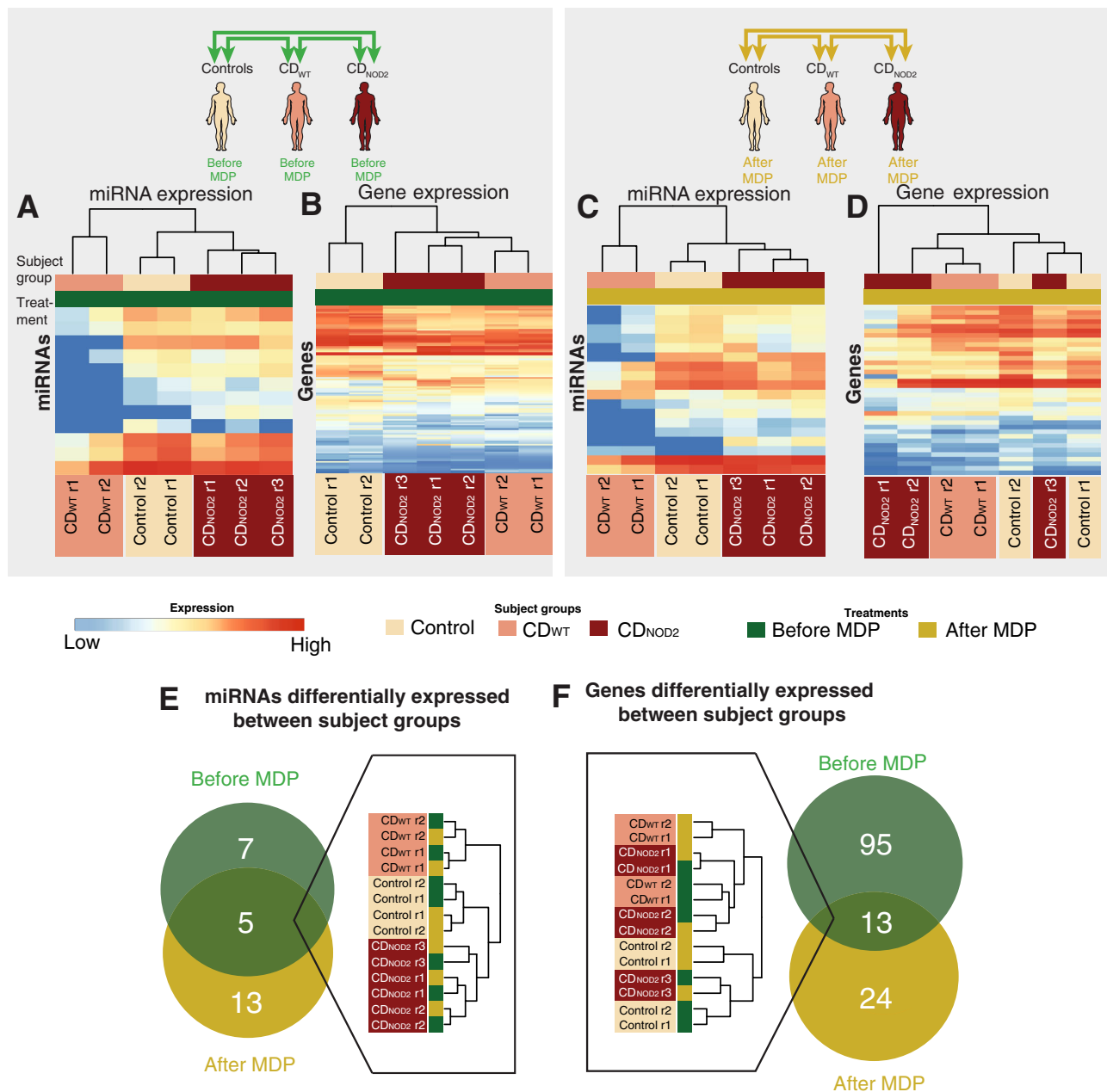
Conversely, the gene expression profiles were drastically different after MDP treatment (Fig. 4c, left panel and Supplementary Fig. 4b) and could distinguish treated vs. non-treated samples with the exception of one  $CD_{NOD2}$  sample (see black arrows in Fig. 4c and Supplementary Fig. 4b). qPCR analysis in the larger validation cohort confirmed the changes in selected differential expressed genes (Fig. 4c, right panel and Supplementary Fig. 4c–f). Notably,  $CD_{WT}$  patients had more than six times the number of upregulated genes than healthy controls (1095 vs. 181, Fig. 4d; also see Supplementary Fig. 4g). Moreover, the large majority of upregulated genes in healthy controls (86%, 156/181) were also upregulated in  $CD_{WT}$  patients. To test whether these 156 shared upregulated genes had the same magnitude of MDP response, we plotted the average gene expression fold change after MDP in the healthy control group vs. corresponding fold change in the  $CD_{WT}$  group (Supplementary Fig. 4h). In general, genes with high upregulation in the healthy control group had even higher upregulation in  $CD_{WT}$  patients. This was also true for the host gene of miR-155, consistent with the higher upregulation of mature miR-155 in  $CD_{WT}$  (Fig. 4a), and indicating an important role for miR-155 in MDP response in CD conditions.

Conversely, in the  $CD_{NOD2}$  patients, no genes were significantly upregulated or downregulated as a response to MDP (only ten genes were induced using the alternative significance calculation method, Supplementary Fig. 4g).

Together, these observations indicated that only subjects with WT *NOD2* (healthy controls and  $CD_{WT}$  patients) had substantial

MDP-induced responses and  $CD_{WT}$  patients had a stronger and wider MDP response. To investigate the responses only found in  $CD_{WT}$  patients further, we mapped the shared upregulated genes between  $CD_{WT}$  and control subjects and the genes only upregulated in  $CD_{WT}$  subjects onto cytokine-receptor KEGG pathways.<sup>34</sup> Strikingly, the receptors for upregulated cytokines were primarily upregulated in  $CD_{WT}$  patients but not in healthy controls (Supplementary Fig. 5a, b), suggesting that the downstream inflammatory response in these pathways was blocked or diminished in healthy controls even if the relevant cytokines were present. These results strongly imply that  $CD_{WT}$  patients had a wider inflammatory response including substantial activation of chemokine receptors and their downstream signaling pathways. Thus, these patients could be characterized as a phenotype with pre-inflammatory hyper-reactive innate response. An investigation of MDP-dependent downregulated genes showed that although the number of downregulated genes was smaller than upregulated genes, the results were consistent to the pattern observed above: on gene level,  $CD_{WT}$  patients had a stronger and wider MDP-mediated inflammatory response than healthy controls, while  $CD_{NOD2}$  patients typically lost the majority of the inflammatory response to MDP (Fig. 4e and Supplementary Fig. 4g, h).

While we observed no significant gene changes upon MDP stimulation in  $CD_{NOD2}$  patients, the MDP states of two  $CD_{NOD2}$  subjects could be distinguished using gene expression profiles (see dendrogram in Fig. 4c, left panel). Thus, we reasoned that subtle expression changes might exist, and that the type of *NOD2* mutation might influence the degree of expression change. To investigate this further, we focused on the genes found

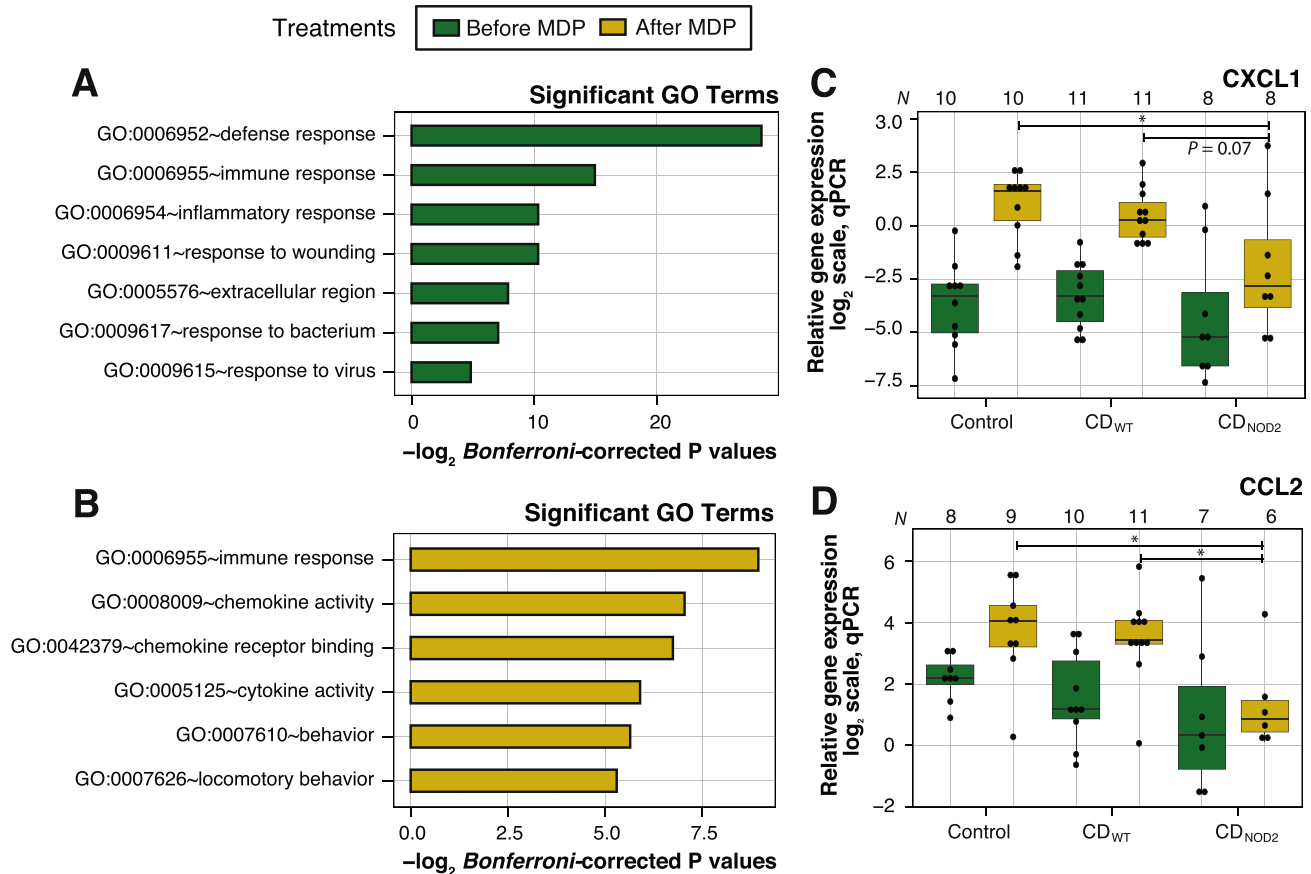


**Fig. 2** miRNA and gene expression patterns between subject groups. **a** Heat map showing the expression pattern of differentially expressed miRNAs between subject groups before MDP stimulation (green arrows in Fig. 1, also shown as schematics on top). Expression values are broken into permilles and assigned to a color gradient ranging from red (high expression) to blue (low expression). Rows indicate individual miRNAs, while columns show individual subjects. Rows and columns are ordered by hierarchical clustering using Euclidian distance, where the top dendrogram indicate similarity between subjects. Subject groups and MDP treatment are indicated by color above the heat maps. Note the clear separation between CD<sub>WT</sub>, CD<sub>NOD2</sub>, and controls. **b** As in **a**, but showing differentially expressed genes (by RNA-Seq) between subject groups before MDP treatment. **c** As in **a**, but showing differentially expressed miRNAs between subject groups after MDP stimulation. Note the clear separation between CD<sub>WT</sub>, CD<sub>NOD2</sub>, and controls. **d** As in **a**, but showing differentially expressed genes between subject groups after MDP treatment. **e** Left: Venn diagram showing the overlap between differentially expressed miRNAs between subject groups before and after MDP treatment. Right: expression-based clustering of the five shared miRNAs as in **a**, but only showing subjects, treatments, and the resulting dendrogram (color codes as in **a–d**). Note that subject groups are clearly separated, while MDP treatment status is not. **f** As in **e**, but for differentially expressed genes. Note that the shared gene set is not as efficient at distinguishing subject groups as the miRNA expression

upregulated and downregulated upon MDP treatment in CD<sub>WT</sub> (1095 and 565 genes, respectively). For each CD<sub>NOD2</sub> sample we plotted the RNA-Seq gene expression fold change after MDP treatment vs. the average gene expression fold change after MDP in the CD<sub>WT</sub> group (Fig. 4f,g, and Supplementary Fig. 4g, right panels). While this confirmed the overall lower inflammatory

response in CD<sub>NOD2</sub> samples vs. CD<sub>WT</sub>, it also showed clear differences in inflammatory response across the individual CD<sub>NOD2</sub> samples. This was consistent between upregulated and downregulated genes and was particularly clear for genes with high fold changes. This indicated the loss of inflammatory responses in CD<sub>NOD2</sub> patients was not binary but a gradient across individuals.





**Fig. 3** Functional annotation of differentially expressed genes between subject groups. **a** Enriched significant GO terms (Bonferroni-corrected  $P$ -values  $\leq 0.05$ ) for differentially expressed genes between subject groups before MDP treatment. X-axis shows  $-\log_2$ (Bonferroni-corrected  $P$ -values) and Y-axis shows GO terms. **b** As in **a**, but showing significant GO terms for differentially expressed genes between subject groups after MDP treatment. **c** qPCR validation for a selected chemokine, CXCL1, which exhibits a large difference between the subject groups after MDP treatment. X-axis shows subject groups before (dark green) and after MDP treatment (yellow). Numbers on top of the plot indicate the number of biological replicates passing the qPCR thresholds. Each subject is visualized as black dots. Y-axis shows the relative gene expression in  $\log_2$  scale. Asterisks indicate level of significance (two-sided  $t$ -test): \* indicates  $P \leq 0.05$ , \*\* $P \leq 0.01$ , and \*\*\* $P \leq 0.001$ . A single comparison with a  $P$ -value close to the significance threshold is indicated. **d** qPCR validation for a selected chemokine, CCL2, following the conventions of **c**

To verify our observation in a larger cohort, we measured the expression of seven validated MDP-responsive genes in CD<sub>WT</sub> (three upregulated genes: CXCL1, CCL2, and IRAK2; four down-regulated genes: MAPK3, MAVS, PALD1, and FLI1) in eight CD<sub>NOD2</sub> and eleven CD<sub>WT</sub> samples before and after MDP treatment by qPCR (Fig. 4h). From this analysis, the large variance in inflammatory response was evident: some CD<sub>NOD2</sub> patients showed virtually no inflammatory response, while some had inflammatory responses comparable to CD<sub>WT</sub> subjects. Overall, in our set, CD<sub>NOD2</sub> patients carrying SNP13 HOM had the lowest responses to inflammation induced by MDP (Fig. 4f–h and Supplementary Fig. 4g).

#### miRNA-gene associations in MDP response

The analyses above did not take miRNA-elicited regulation of gene expression into account. To identify likely regulatory miRNA-gene associations, we used experimentally verified miRNA targets from mirTarBase 4.3<sup>35</sup> and required that the miRNA and gene pair in question had anti-correlated expression patterns (e.g., upregulation of miRNA and downregulation of gene, or vice versa) in any of the previously made pairwise expression analyses (any arrow in Fig. 1). When including all the differentially expressed miRNAs and genes, only 16 putative miRNA-gene pairs centered on miR-155 were identified (Fig. 5a). To further consolidate the regulatory role of miR-155, we transfected pre-miR-155 into human monocytic

THP-1 cells and measured the expression change of the 16 candidate target genes with qPCR (Fig. 5a). Of these, PALD1, FLI1, PCYOX1, and BRI3BP were significantly repressed upon miR-155 transfection ( $P \leq 0.05$ , two-sided  $t$ -test, Fig. 5b–e), while TYSND1, TM6SF1, and SKI were borderline significant (Fig. 5f–h, for details see “Methods”). It is important to consider that this type of experiment may include indirect effects; however, several of the genes had additional external evidence in the form of gene reporter assays (external evidence is summarized in Fig. 5i).

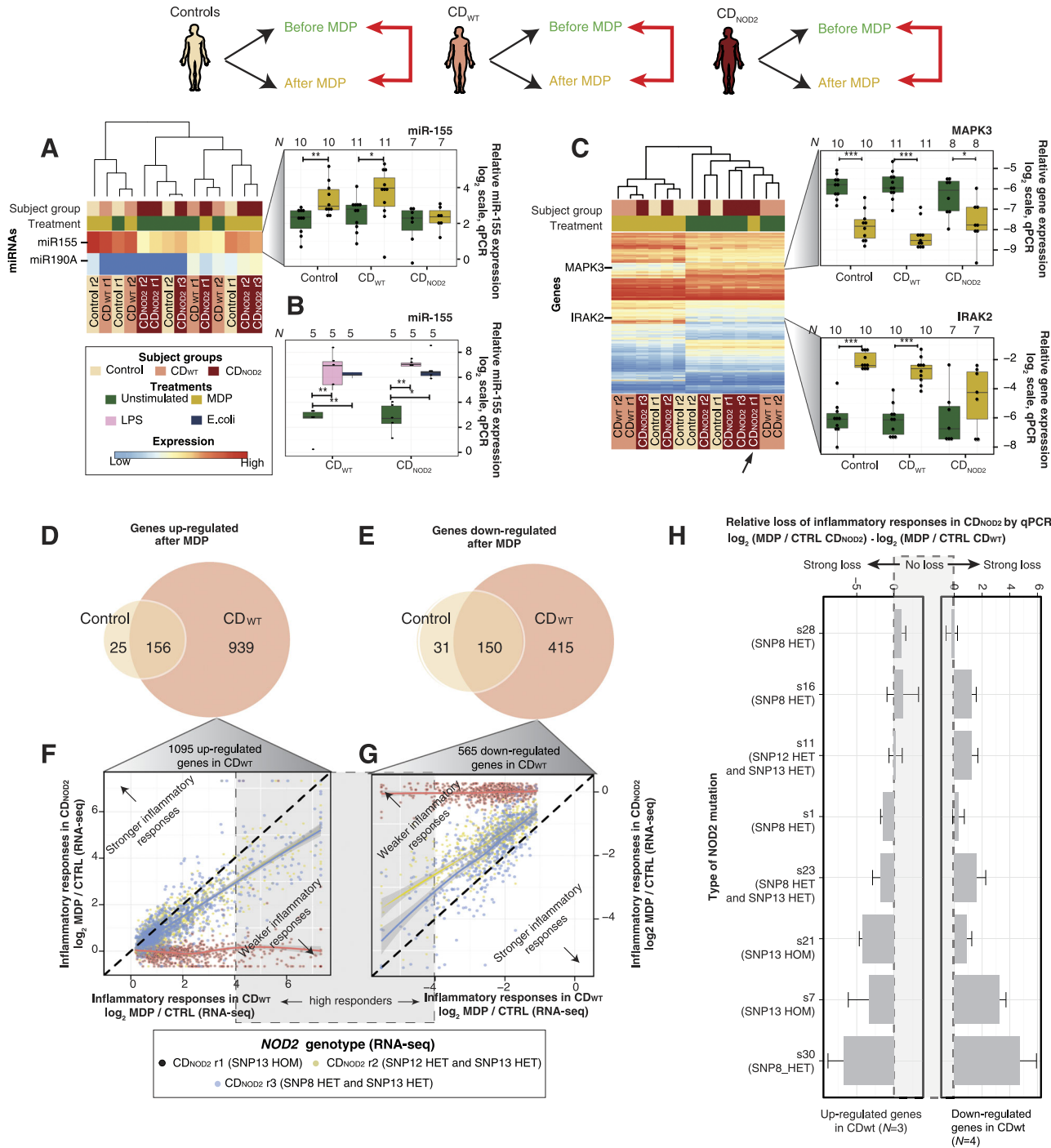
We found that this limited set of seven validated miR-155 targeting genes alone could distinguish MDP-treated cells from non-treated cells except for the same CD<sub>NOD2</sub> subject that could not be correctly classified using the whole gene set (Fig. 6a). Because we identified a large variance across CD<sub>NOD2</sub> patients on gene expression level following MDP stimulation, we hypothesized that a similar variation existed on miR-155 regulation level. In order to test this, we focused on miR-155 and the seven target genes identified above, and compared their MDP response between individual CD<sub>NOD2</sub> patients and pooled CD<sub>WT</sub> patients (Supplementary Fig. 6), similarly to our previous gene expression comparison in Fig. 4f,g. As expected, the CD<sub>NOD2</sub> subject that showed the smallest inflammatory responses at gene level also displayed the most drastic loss of miR-155 upregulation and downregulation of its seven validated targets (Supplementary Fig. 6). This variance across patients was confirmed by qPCR analysis in

the larger cohort. This showed that the loss of miR-155 induction following MDP treatment varied across subjects, and the corresponding level of downregulation of the miR-155 target genes generally followed the miR-155 levels (Fig. 6b). In general, CD<sub>NOD2</sub> subjects displaying a reduced MDP response of gene level also showed this on miR-155 level (Figs. 4h, 6b).

DISCUSSION

In this study we explored gene and miRNA expression in an initial set of human blood monocytes using genome-wide high-throughput sequencing techniques and verified our major

conclusions by applying qPCR analyses in a larger validation cohort or THP-1 cells. All the samples were split by MDP treatment and were broken up by subject groups, including patients with NOD2 genotypes. The results from pairwise differential expression studies between the subjects demonstrated that the MDP treatment resulted in considerable changes in gene expression. In particular, CD<sub>WT</sub> patients showed a much stronger and wider inflammation response compared to healthy controls. This is in line with a report investigating the regulation of specific inflammatory pathways upon innate engagement of monocytes in CD,<sup>36</sup> but we here show that these changes are more general, involving activating



**Fig. 4** MDP-dependent miRNA and gene expression response within subject groups. **a Left:** Heat map as in Fig. 2a showing the two miRNAs that were differential expressed as a response to MDP within a subject group (schematic on top shows the analyzed comparisons). **Right:** qPCR validations for miR-155 are following the conventions of Fig. 3c. Asterisks indicate level of significance (two-sided paired *t*-test): \* indicates  $P \leq 0.05$ , \*\* $P \leq 0.01$ , and \*\*\* $P \leq 0.001$ . **b** qPCR analysis of miRNA-155 expression LPS or intact *E. coli* in monocytes from CD patients, broken up by *NOD2* genotype and treatment. Y-axis shows the  $\log_2$  normalized fold change of miR-155 vs. expression control (see Supplementary Methods for normalization); dots stand for individual patients. Asterisks indicate level of significance (two-sided paired *t*-test): \* indicates  $P \leq 0.05$ , \*\* $P \leq 0.01$ , and \*\*\* $P \leq 0.001$ . **c Left:** As in left of **a**, but showing genes that were differential expressed as a response to MDP within a subject group. The black arrow highlights the outlier CD<sub>NOD2</sub> subject that shows no substantial change in response to MDP. **Right:** qPCR validations of *IRAK2* and *MAPK3* genes are shown in the right panel, organized as in the right of panel **a**. **d** Venn diagrams showing the overlap of upregulated genes in response to MDP within subject groups. **e** Venn diagrams showing the overlap of downregulated genes in response to MDP within subject groups. **f** Degree of inflammatory gene response of individual CD<sub>NOD2</sub> patients compared to the average response of CD<sub>WT</sub> patients from the exploratory cohort. The inflammatory response is measured in a set of 1095 genes that were upregulated in CD<sub>WT</sub>. X-axis shows  $\log_2$  fold change of the 1095 genes in CD<sub>WT</sub> patients after MDP. Y-axis shows corresponding fold changes in individual CD<sub>NOD2</sub> patients. Each dot corresponds to one gene, colored by individual CD<sub>NOD2</sub> patients. Solid lines represent fitted linear models for each CD<sub>NOD2</sub> patient using the loess approach; 95% confidence intervals are indicated as gray shadows. The dashed diagonal in black corresponds to equivalent inflammatory response between CD<sub>NOD2</sub> and CD<sub>WT</sub> patients. Genes located above the diagonal will correspond to stronger inflammatory response in a CD<sub>NOD2</sub> patient than the average CD<sub>WT</sub> response. Conversely, genes located below the diagonal will correspond to decreased inflammatory response in CD<sub>NOD2</sub> vs. CD<sub>WT</sub> patients. Genes with high  $\log_2$  fold changes ( $>4$ ) are considered high MDP responders (see main text). **g** As in **e**, but for the 565 downregulated genes in CD<sub>WT</sub> patients following MDP stimulation. The two parts of the plot indicating stronger or weaker inflammatory response in a CD<sub>NOD2</sub> are now swapped to the opposite. Genes with low  $\log_2$  fold changes ( $<-4$ ) are considered high MDP responders (see main text). **h** Relative loss of inflammatory response in individual CD<sub>NOD2</sub> patients compared to average CD<sub>WT</sub> patients in the validation cohort by qPCR. Each pair of horizontal bar plots shows the average degree of inflammatory response loss for a CD<sub>NOD2</sub> patient, expressed as the average difference vs. CD<sub>WT</sub> patients for three upregulated genes in CD<sub>WT</sub> patients (*CXCL1*, *CCL2*, and *IRAK2*, left panel) and four downregulated genes in CD<sub>WT</sub> patients (*MAPK3*, *PALD1*, *MAVS*, and *FLI1*, right panel). Error bars indicate the standard error of the mean

cytokine receptors and their downstream signaling pathways. In addition, it is noteworthy that while the major inflammatory responses in the gene layer of CD<sub>WT</sub> patients corresponded to upregulation, miRNA expression was generally downregulated in these patients. This suggests that the innate hyper-responses upon inflammation in CD<sub>WT</sub> patients might be also mediated by generally reduced activity of the miRNA regulatory layer. Such a *NOD2*-dependent phenotype of a pre-inflammatory innate hyper-responsiveness state could be essential in shaping the inflammatory processes found in CD. In line with this, there is evidence of an increased risk of onset of CD in patients after infections with intestinal pathogens, e.g., *Salmonella* spp. or *Campylobacter* spp.<sup>37, 38</sup>

In contrast to mRNA levels, miRNA expression signatures showed large differences between subject groups, but only two miRNAs changed significantly with MDP treatment including a key inflammatory regulator, miR-155. Studies investigating *NOD2*-specific miR-155 inductions have shown MDP stimuli alone has limited impact on miR-155 expression in human dendritic cells,<sup>15</sup> mouse macrophages,<sup>27</sup> and embryonic fibroblasts.<sup>39</sup> Our results demonstrated that MDP could induce miR-155 and its host gene in human monocytes expressing WT *NOD2*, and this induction was particularly strong in monocytes from CD patients, suggesting the *NOD2*-based induction on miR-155 might differ across both cell and disease types.

Previous studies have shown that the *NOD2* loss-of-function SNPs lead to decreased immune response,<sup>40, 41</sup> but the differences between individual subjects and type of mutations were not addressed. Here, we showed that the effect of *NOD2* loss-of-function SNPs is diverse: some individuals from our CD<sub>NOD2</sub> cohort lost the MDP response, while some could mount an inflammatory response similar to that of WT *NOD2* subjects. This observation was reflected on both general gene expression profiles, on miR-155 levels and on its target genes. Patients carrying the SNP13 HOM appeared to have the largest response loss, although to investigate the detailed effect of respective genotype, expression quantitative trait loci analyses across a much larger population will be necessary.

As we have not induced or rescued *NOD2* mutations, these results are based on correlations between genotype and phenotype. However, as reduced induction of miR-155 in the CD<sub>NOD2</sub> group occurred only when treating with MDP (a *NOD2*

agonist) and not when treating with LPS or intact bacteria, it is highly likely that the *NOD2* mutations are the main drivers for the phenotypes observed and by extension, that inflammatory pathways not involving *NOD2* are likely intact in these patients. Additional mutations carried by these patients may modulate the MDP response and may explain the wide MDP-response repertoire in CD<sub>NOD2</sub> patients. Such mutations may not only affect protein-coding genes but also regulatory sites, including the disruption of miRNA-binding sites, which have been reported to occur in immune-associated genes, including *NOD2* and *IRGM*.<sup>42, 43</sup> However, to analyze such cases, genome-wide genotyping is necessary.

In summary, our study identified substantial and novel changes in *NOD2* signaling of CD<sub>NOD2</sub> and CD<sub>WT</sub> patients. The first group had on average dampened MDP responses, but the global response was highly varied across individuals, on both gene and miR-155 regulation levels. The CD<sub>WT</sub> group was characterized by a pre-inflammatory innate hyper-responsive phenotype and thus responds much stronger to MDP than cells from healthy controls. This hyper-responsive phenotype was characterized by activation a wide range of cytokine receptors and potentially also by a general downregulation of miRNA expression in these patients. Both types of dysregulated innate responses could, in different ways, increase susceptibility to intestinal inflammation in response to commensal bacteria. Moreover, we have shed light on an underappreciated role of miRNAs in identifying disease-specific and genotype-specific alterations in CD.

## METHODS

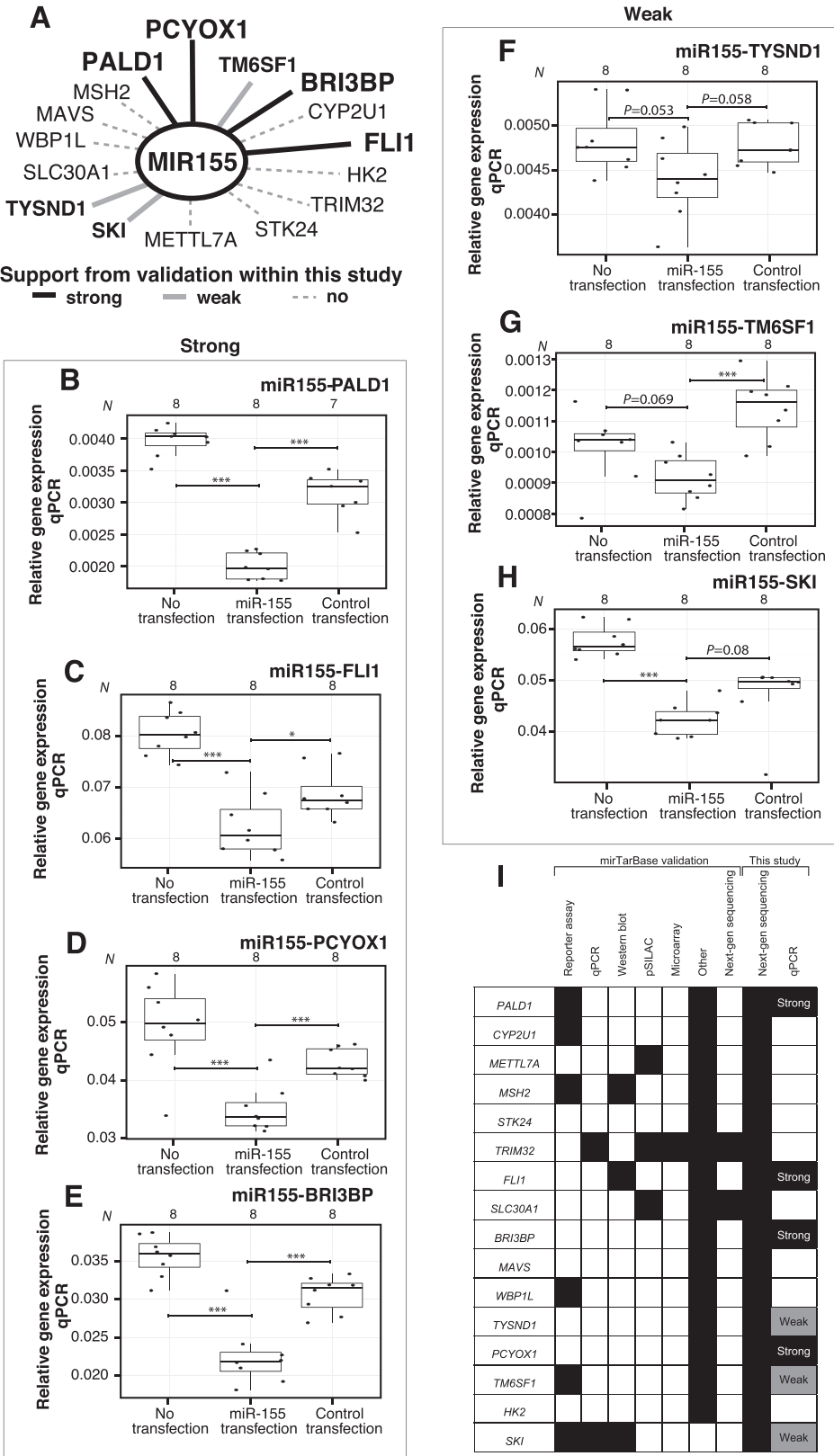
### Study population and genotyping

The patient population in this study was derived from a total of 57 healthy controls (i.e., without any known GI disorders) and 236 patients with well-established CD attending the IBD Clinic at Department of Gastroenterology, Medical Section, Herlev Hospital. The diagnostic criteria for CD were based on established guidelines<sup>44</sup> applying clinical, radiologic, endoscopic, and histologic criteria. Screening for the common *NOD2* polymorphisms (i.e., SNP8 [Arg702Trp], SNP12 [Gly908Arg], and SNP13 [Leu1007fsinsC]) was performed as previously described.<sup>45</sup> All participants in this study were WT in respect to *autophagy-related 16-like 1* (*ATG16L1*) genotype. Patients included in the sequencing and validation studies were in complete clinical remission for at least 4 weeks prior to inclusion (Harvey-Bradshaw index  $\leq 5$ ),<sup>46</sup> and did not receive glucocorticoids or biologics for at least 3

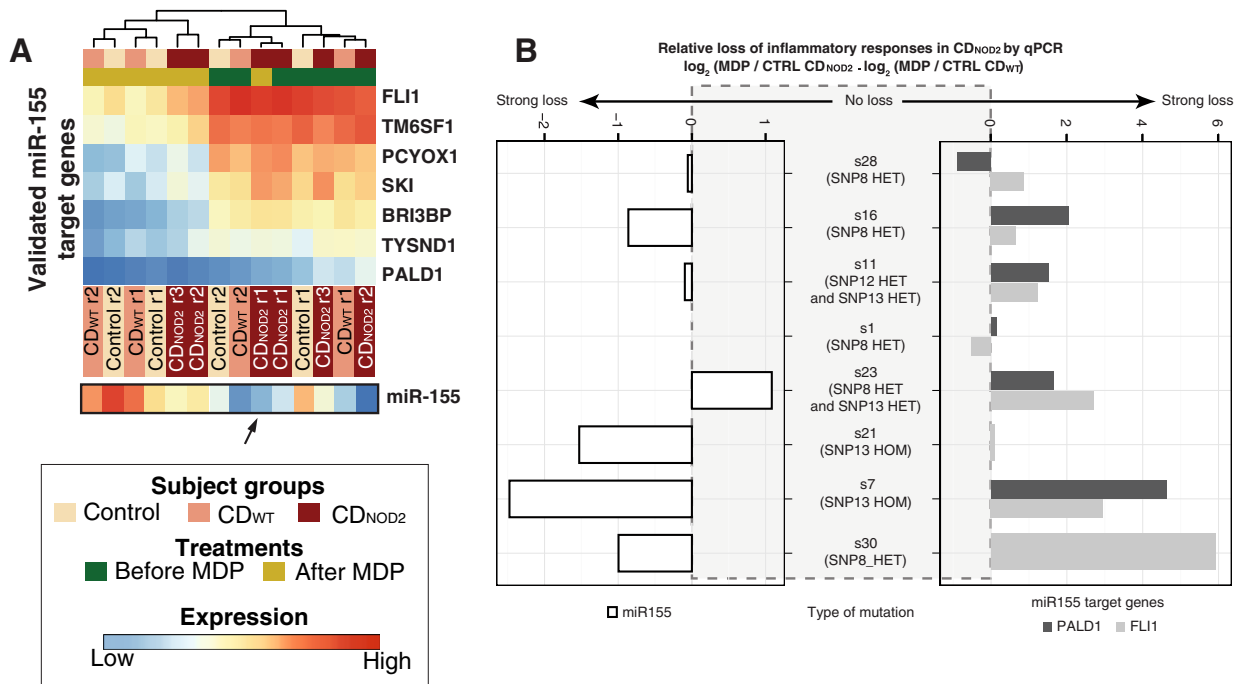


8 preceding months. Long-term treatment with thiopurines was, however, allowed if the dosing had been stable for more than 2 months. For RNA-Seq and small RNA-Seq, monocytes were isolated from peripheral blood samples from three subject groups ( $N = 7$  in total): (I) healthy controls with WT *NOD2* ( $N = 2$ ), (II) CD patients with WT *NOD2* ( $CD_{WT}$ ) ( $N = 2$ ), and (III) CD

patients with *NOD2* SNPs ( $CD_{NOD2}$ ; one SNP13 homozygote—replicate 1, one SNP13 heterozygote and SNP12 heterozygote—replicate 2, and one SNP13 heterozygote and SNP8 heterozygote—replicate 3) ( $N = 3$ ). Subsequently, miRNA/mRNA patterns identified in the RNA-Seq exploratory experiments were tested in a larger cohort ( $N = 29$  in total) where



**Fig. 5** Identification of miR-155-gene associations. **a** miRNA-gene association network showing miR-155-gene interactions identified. Line color summarizes the qPCR-based transfection analysis results in the exploratory cohort (shown in **b–h**). Strong validations (**bold** lines, corresponding to **b–e**) and weak validations (**gray** lines, corresponding to **f–h**), and interactions that failed validations are shown as **dotted** lines. Strong and weak criteria are described in “Methods” (**b–h**). qPCR results of miR-155 transfected THP1 cells. Y-axis shows relative qPCR expression of the gene indicated on **top** of the plot. X-axis shows boxplots for cells that were not transfected, miR-155-transfected cells, and control-transfected cells. **Dots** indicate individual samples. **Asterisks** indicate level of significance (two-sided *t*-test): \* indicates  $P \leq 0.05$ , \*\* $P \leq 0.01$ , and \*\*\* $P \leq 0.001$ . *P*-values for tests that are borderline significant are shown (see Methods). **b–e** show strong validation cases, **f–h** shows weak validation cases. **i** Summary of evidence for miR-155 regulation of the selected genes, from mirTarBase and this study. **Black** boxes indicate the evidence from the indicated type of experiments from mirTarBase. “Strong” and “weak” categories refer to the classification of validations in **b–h**



**Fig. 6** Classification power and variability of miR-155 and its targets across CD<sub>NOD2</sub> patients. **a** Heat map as in Fig. 2a showing the expression pattern of the seven genes interacting with miR-155 indicated in **a**. The expression of miR-155 in corresponding libraries is indicated below the heat map. The **black** arrow highlights the outlier CD<sub>NOD2</sub> subject that shows no substantial response to MDP. **b** Degree of response changes of individual CD<sub>NOD2</sub> patients compared to the average response of CD<sub>WT</sub> patients from the exploratory cohort on miR-155 level (**left** panel) and two selected miR-155 target genes (PALD and FLI1, **right** panel). The plot follows the same conventions as Fig. 4h

monocytes were obtained from (I) healthy controls with WT *NOD2* ( $N = 10$ ), (II) CD<sub>WT</sub> ( $N = 11$ ), (III) CD<sub>NOD2</sub> (mixed *NOD2* polymorphisms) ( $N = 8$ ). The monocytes isolated from each participant were grown with or without MDP stimulation (see Supplementary Methods) for sequencing and qPCR analysis. Supplementary Table 1 shows additional subject data. Validations of miRNA-gene interactions were done with eight biological replicates in THP-1 cells (see Supplementary Materials and Methods).

#### Real-time qPCR validation of mRNAs and miRNAs

RNA was purified from monocytes and THP-1 cells, and converted to cDNA as described in Supplementary Methods. Primer sequences are shown in Supplementary Table 2. For each primer pair, only the samples that passed qPCR thresholds were used in respective analyses (see “Methods”). Numbers of samples passing thresholds are shown on top of corresponding figures. Two-sided *t*-tests qPCR values were used as default statistical tests between two groups. Paired tests were conducted for comparisons within subject groups shown in Fig. 4 and Supplementary Fig. 4.  $P \leq 0.05$  was used as significance cutoff. For miRNA transfection experiments in THP-1 cells, validations were considered “strong” if miR-155-transfected samples vs. no transfected samples and miR-155-transfected samples vs. control-transfected samples were both significant. Validations were considered “weak” either if  $P < 0.06$  in both comparisons, or if  $P \leq 0.05$  in one of the two comparisons. Validations were considered “negative” if none of the above applied.

#### RNA-Seq and small RNA-Seq preparation, mapping and quantification

Total RNA and short RNAs were purified using the Nucleospin RNA II isolation kit (Macherey-Nagel) as recommended by the manufacturer. Quantity and purity were determined on a NanoDrop1000, all samples had a 260/280 nm absorbance  $> 1.9$ . Illumina libraries were constructed using TruSeq Stranded mRNA Sample Prep Kit and TruSeq Small RNA Library Preparation Kit (Illumina) following the manufacturer’s protocols. The libraries were sequenced on an Illumina Genome Analyser IIx running sequencing 76 cycles, single read, for the mRNA libraries, and 36 cycles, single read, for the short RNA libraries. Adapters were trimmed from reads, which were then, filtered by quality and mapped to the hg19 assembly (see Supplementary Methods).

#### Differential expression analysis

Since our samples were derived from three subject groups and two MDP conditions, many types of pair-wise comparisons between groups were feasible, which would answer different types of questions (Fig. 1). Three group comparisons were performed as shown in Fig. 1: (1) between patient groups “before MDP treatment” (green arrows), (2) between subject groups “after MDP treatment” (yellow arrows), and (3) between expression profiles of the same subject group before and after treatment “MDP response” (red arrows).

For differential expression of genes in the above pair-wise comparisons, we conducted a combined analysis using DESeq<sup>47</sup> and Cuffdiff<sup>48</sup>, where we required an adjusted  $P$ -value  $\leq 0.05$  (Benjamini–Hochberg False Discovery Rate, FDR) from DESeq combined with a non-adjusted  $P$ -value  $\leq 0.05$  from Cuffdiff (shown in main figures). The reciprocal analysis requiring FDR  $\leq 0.05$  from Cuffdiff combined with a non-adjusted  $P$ -value  $\leq 0.05$  from DESeq did not give substantially different results (shown in Supplementary Figs. 1, 3–5). Differential expression analysis for miRNAs was performed by DESeq<sup>47</sup> requiring a FDR  $\leq 0.05$ . All differentially expressed genes analyzed with both methods and miRNAs are listed in Supplementary Table 3.

### Code availability

Aside from the programs mentioned above, computational analyses were performed in R. Code is available by request.

### Ethical approval

The study was approved by the Scientific Ethics Committee of the Copenhagen Capital Region. Written informed consents from all included subjects were obtained prior to participation, and the project fulfilled the Helsinki V Declaration.

### ACKNOWLEDGEMENTS

This work was supported by grants from Axel Muusfeldts Foundation, Holger og Ruth Hesses Mindefond, Else og Mogens Wedell-Wedellsborgs Foundation, Direktør Kurt Bønnelycke og hustru fru Grethe Bønnelyckes Foundation, P. A. Messerschmidt og Hustrus Foundation, the Foundation of Aase and Ejnar Danielsen, the Danish Council for Independent Research, the Lundbeck Foundation, and the Novo Nordisk Foundation. We thank Hanne Fuglsang, Nine Scherling, Noor Irmam, and Tine Helene Soendergaard for excellent technical assistance. We thank Kim Jensen for access to PCR instruments.

### DATA SHARING

Small RNA and RNA-Seq data are deposited in the GEO database, ID: GSE69446.

### AUTHOR CONTRIBUTIONS

Y.C. and M.S. are shared first authors. M.B. and J.B. are shared second authors. A.S., J.B. S., and O.H.N. designed the study and supervised the project. Y.C. analyzed the data. Y.C. and A.S. made the figures. M.S. made the sequencing libraries. M.S., M.B., J.B., Y.L., and M.C. made qPCR and miRNA validations. J.B.S. and O.H.N. did patient allocation. Y.C., M.S., J.B.S., O.H.N., and A.S. wrote the manuscript.

### COMPETING INTERESTS

The authors declare that they have no competing interests.

### REFERENCES

- Baumgart, D. C. & Sandborn, W. J. Crohn's disease. *Lancet* **380**, 1590–1605 (2012).
- Danese, S. & Fiocchi, C. Ulcerative colitis. *N. Engl. J. Med.* **365**, 1713–1725 (2011).
- Abraham, C. & Cho, J. H. Inflammatory bowel disease. *N. Engl. J. Med.* **361**, 2066–2078 (2009).
- Kalla, R., Ventham, N. T., Satsangi, J. & Arnott, I. D. R. Crohn's disease. *Br. Med. J.* **349**, g6670–g6670 (2014).
- Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
- McGovern, D. P. B., Kugathasan, S. & Cho, J. H. Genetics of inflammatory bowel diseases. *Gastroenterology* **149**, 1163–1176.e2 (2015).
- Braat, H., Peppelenbosch, M. P. & Hommes, D. W. Immunology of Crohn's disease. *Ann. N. Y. Acad. Sci.* **1072**, 135–154 (2006).
- Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
- Ogura, Y. *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
- Mo, J. *et al.* Pathogen sensing by nucleotide-binding oligomerization domain-containing protein 2 (NOD2) is mediated by direct binding to muramyl dipeptide and ATP. *J. Biol. Chem.* **287**, 23057–23067 (2012).
- Salem, M., Seidelin, J. B., Rogler, G. & Nielsen, O. H. Muramyl dipeptide responsive pathways in Crohn's disease: from NOD2 and beyond. *Cell. Mol. Life Sci.* **70**, 3391–3404 (2013).
- Watanabe, T. *et al.* Muramyl dipeptide activation of nucleotide-binding oligomerization domain 2 protects mice from experimental colitis. *J. Clin. Invest.* **118**, 545–559 (2008).
- Seidelin, J. B., Broom, O. J., Olsen, J. & Nielsen, O. H. Evidence for impaired CARD15 signalling in Crohn's disease without disease linked variants. *PLoS ONE* **4**, e7794 (2009).
- Chen, Y. *et al.* miR-122 targets NOD2 to decrease intestinal epithelial cell injury in Crohn's disease. *Biochem. Biophys. Res. Commun.* **438**, 133–139 (2013).
- Brain, O. *et al.* The intracellular sensor NOD2 induces microRNA-29 expression in human dendritic cells to limit IL-23 release. *Immunity* **39**, 521–536 (2013).
- Chuang, A. Y., Chuang, J. C., Zhai, Z., Wu, F. & Kwon, J. H. NOD2 expression is regulated by microRNAs in colonic epithelial HCT116 cells. *Inflamm. Bowel. Dis.* **20**, 126–135 (2014).
- Kalla, R. *et al.* microRNAs: new players in IBD. *Gut* **64**, 504–517 (2015).
- Sayed, D. & Abdellatif, M. microRNAs in development and disease. *Physiol. Rev.* **91**, 827–887 (2011).
- Bai, J. *et al.* Integrating analysis reveals microRNA-mediated pathway crosstalk among Crohn's disease, ulcerative colitis and colorectal cancer. *Mol. Biosyst.* **10**, 2317–2328 (2014).
- O'Connell, R. M., Taganov, K. D., Boldin, M. P., Cheng, G. & Baltimore, D. microRNA-155 is induced during the macrophage inflammatory response. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1604–1609 (2007).
- Jablonski, K. A., Gaudet, A. D., Amici, S. A., Popovich, P. G. & Guerau-de-Arellano, M. Control of the inflammatory macrophage transcriptional signature by miR-155. *PLoS ONE* **11**, e0159724 (2016).
- Seddiki, N., Brezar, V., Ruffin, N., Lévy, Y. & Swaminathan, S. Role of miR-155 in the regulation of lymphocyte immune function and disease. *Immunology* **142**, 32–38 (2014).
- Van der Goten, J. *et al.* Integrated miRNA and mRNA expression profiling in inflamed colon of patients with ulcerative colitis. *PLoS ONE* **9**, e116117–e116117 (2013).
- Fasseu, M. *et al.* Identification of restricted subsets of mature microRNA abnormally expressed in inactive colonic mucosa of patients with inflammatory bowel disease. *PLoS ONE* **5**, e13160 (2010).
- Singh, U. P. *et al.* miR-155 deficiency protects mice from experimental colitis by reducing T helper type 1/type 17 responses. *Immunology* **143**, 478–489 (2014).
- Wang, P. *et al.* Inducible microRNA-155 feedback promotes type I IFN signaling in antiviral innate immunity by targeting suppressor of cytokine signaling 1. *J. Immunol.* **185**, 6226–6233 (2010).
- Koch, M., Mollenkopf, H.-J., Klemm, U. & Meyer, T. F. Induction of microRNA-155 is TLR- and type IV secretion system-dependent in macrophages and inhibits DNA-damage induced apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1153–E1162 (2012).
- Tili, E. *et al.* Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock. *J. Immunol.* **179**, 5082–5089 (2007).
- Economou, M., Trikalinos, T. A., Loizou, K. T., Tsianos, E. V. & Ioannidis, J. P. A. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am. J. Gastroenterol.* **99**, 2393–2404 (2004).
- Strober, W., Asano, N., Fuss, I., Kitani, A. & Watanabe, T. Cellular and molecular mechanisms underlying NOD2 risk-associated polymorphisms in Crohn's disease. *Immunol. Rev.* **260**, 249–260 (2014).
- Pathak, S. *et al.* miR-155 modulates the inflammatory phenotype of intestinal myofibroblasts by targeting SOCS1 in ulcerative colitis. *Exp. Mol. Med.* **47**, e164 (2015).
- Xu, G. *et al.* Transcriptome and targetome analysis in MIR155 expressing cells using RNA-seq. *RNA* **16**, 1610–1622 (2010).
- O'Connell, R. M., Rao, D. S. & Baltimore, D. microRNA regulation of inflammatory responses. *Annu. Rev. Immunol.* **30**, 295–312 (2012).
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
- Hsu, S.-D. *et al.* miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res.* **42**, D78–D85 (2013).
- Santaolalla, R., Fukata, M. & Abreu, M. T. Innate immunity in the small intestine. *Curr. Opin. Gastroenterol.* **27**, 125–131 (2011).

37. Gradel, K. O. *et al.* Increased short- and long-term risk of inflammatory bowel disease after salmonella or campylobacter gastroenteritis. *Gastroenterology* **137**, 495–501 (2009).
38. García Rodríguez, L. A., Ruigómez, A. & Panés, J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterology* **130**, 1588–1594 (2006).
39. Schulte, L. N., Westermann, A. J. & Vogel, J. Differential activation and functional specialization of miR-146 and miR-155 in innate immune sensing. *Nucleic Acids Res.* **41**, 542–553 (2013).
40. Li, J. *et al.* Regulation of IL-8 and IL-1 $\beta$  expression in Crohn's disease associated NOD2/CARD15 mutations. *Hum. Mol. Genet.* **13**, 1715–1725 (2004).
41. Moreira, L. O. *et al.* The TLR2-MyD88-NOD2-RIPK2 signalling axis regulates a balanced pro-inflammatory and IL-10-mediated anti-inflammatory cytokine response to gram-positive cell walls. *Cell. Microbiol.* **10**, 2067–2077 (2008).
42. Landi, D. *et al.* Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. *Carcinogenesis* **29**, 579–584 (2008).
43. Brest, P. *et al.* A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat. Genet.* **43**, 242–245 (2011).
44. Van Assche, G. *et al.* The second European evidence-based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis. *J. Crohns Colitis* **4**, 7–27 (2010).
45. Yazdanyar, S. & Nordestgaard, B. G. NOD2/CARD15 genotype and common gastrointestinal diseases in 43,600 individuals. *J. Intern. Med.* **267**, 228–236 (2010).
46. Harvey, R. F. & Bradshaw, J. M. A simple index of Crohn's-disease activity. *Lancet* **1**, 514 (1980).
47. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
48. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578, doi:10.1038/nprot.2012.016 (2012).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *npj Genomic Medicine* website (doi:10.1038/s41525-016-0001-4).